ESTIMATION OF GENETIC GAIN IN MILK YIELD BY SELECTION IN A CLOSED HERD OF DAIRY CATTLE

By J. M. RENDEL AND ALAN ROBERTSON

Animal Breeding and Genetics Research Organization,
Institute of Animal Genetics, Edinburgh

(With One Text-figure)

INTRODUCTION

The simplest and most direct way of breeding farm animals is by mass selection, that is, by choosing breeding stock purely on the basis of their own phenotype and mating the chosen animals at random. In dairy cattle the value of this practice is reduced, since such selection cannot be practised on males. It has therefore been suggested that the bull should be judged on the performance of his progeny in order to make up for the fact that he has no measurable phenotype himself. In this way progress would be hastened. The advocates of this view can point to evidence in historical documents and claim that the cattle of centuries ago were not greatly different from our own in milk yield, which argues the inefficiency of any mass selection which may have taken place since then. However, such evidence, though interesting historically, does not carry much scientific weight; and it seems desirable to examine the method of mass selection, first, from the theoretical point of view in order to estimate its effect under the most favourable circumstances and, secondly, from the practical aspect in considering how much selection is actually practised in herds and how much genetic gain can be expected therefrom.

PREVIOUS ESTIMATES OF GENETIC GAIN FOR DAIRY HERDS

It is obvious that the direct observation of the variation of the average yield of a herd gives little indication of the genetic improvement achieved. There are many reasons why the yield of a herd should increase with time—the owner will increase his skill at managing his herd with practice, and there has been a great improvement in methods of feeding dairy cattle since the 1920's. In the reverse direction, there is the influence of two wars to be reckoned with, and the recovery from these provides two further periods of general increase in milk yields.

At first sight it would seem possible to measure the effect of such extraneous changes on yields by considering the yield of the same cow in successive years. After subtracting environmental trends so determined from the total, the genetic changes are left. This method is used by Lörtscher (1937) and with slightly more sophisticated statistics by Nelson (1943). There is, however, a logical error in this approach which vitiates the conclusions drawn. The records of the same cows in successive years will vary systematically for two reasons: first, changes in management whose effect we wish to evaluate, and secondly, increase in age of the cow. Before comparison of records in successive years is made, therefore, all yields must be reduced to their equivalent at a standard age. To do this, correction factors must be used; if these are calculated from the data provided by the herd under review, they will include the changes due to either environmental or genetic

trends. Suppose, for example, feeding is being steadily improved and yields are rising for this reason; suppose also that yields are rising because of selection leading to genetic gains; we shall then have the picture given in Fig. 1.

In the figure, the position is simplified by considering only two lactations. Point A is the level of first lactations in year 0, and B is the level of second lactations of the same cows in year 1. The genetic and management levels are assumed to be improving at the rate of g and m per year. BE is the apparent age increase, which is, in fact, made up of the true age effect a and the management improvement m. The first lactations in year 1 will be at the point C. We can actually measure the increase in yield of a given cow with time, $a + m$, and the increase in the mean yield of the herd, $g + m$. If we calculate the
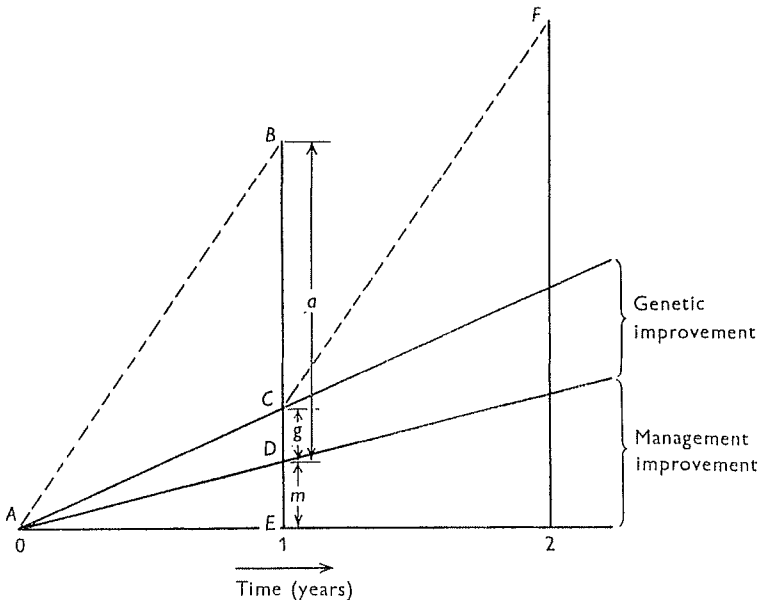

Fig. 1.

age correction factor from $a + m$, assuming it is equal to a, then when the conversion to standard yield is made both m and a are taken out, leaving no apparent effect of environment. If, on the other hand, the factor is calculated by comparing first and second lactations made in the same year a slope corresponding to $a - g$ is being measured, and the correction is then reduced by the amount of any genetic trend, so that the environmental effect is made to include the genetic. In this case, the effects of selection must be allowed for. Algebraically we can say that $g + m$ and $m + a$ can be measured, from which we can deduce $a - g$, which can also be estimated without first knowing $g + m$ and $m + a$, but none of the three entities can be known separately.

Correction factors may be taken from estimates made in other herds. These will have against them the objection already outlined and also doubt as to their suitability. The effect of age on yield is not independent of management. A cow which is matured early by heavy feeding will not show as marked an age effect as a cow reared on a lower plane of nutrition (Bonnier, Hansson & Skjervold, 1948), so that correction factors must be fitted to the management of the herd in question if they are to be appropriate. If, for instance, we have a quick-maturing herd, either for genetic or management reasons, use of corrections

derived from slower-maturing animals in this method will suggest that management is getting worse (cows' corrected first records will be above the true value), and consequently the genetic level improving, even though there is really no change in either. It may be said that such errors will be small, but the age effect is large compared with the other effects, and any errors in its estimation will be cumulative. Suppose the factor converting first to second lactation is 0·005 too high, then first lactations when converted will be 0·5 % higher than second (other things being equal), the environment will appear to be deteriorating at 0·5 % per year, and the genetic level will be correspondingly increased. Since the actual trends are of the order of 0·75 % per year, errors of this size are of some importance, and are likely to be exceeded in most estimates in practice.
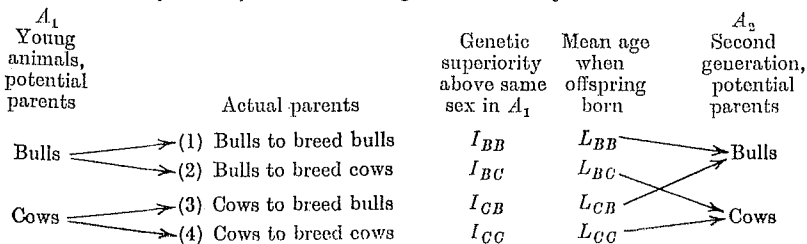
It must be concluded that such indirect estimates of the genetic gains under selection are useless. The alternative is to estimate the genetic gains directly from the selection applied. This is the method used here.

IMPROVEMENT TO BE EXPECTED FROM DIRECT SELECTION UNDER OPTIMUM CONDITIONS

The application of the principles of genetics to the problems arising in breeding farm animals for economical production are of comparatively recent origin. They are based on the theoretical considerations of Wright, Fisher and Haldane. The detailed application of their findings to animal breeding has been mostly due to Lush and his co-workers. The general principles of this new approach are given by Lush in his book *Animal Breeding Plans*. The basic concept, that of 'heritability', is a statistical one, relating the phenotype of an animal to its genotypic value. In its narrow sense, as we shall apply it here to individuals, it can be defined as the regression of genotypic value on phenotype. The heritability of milk yield based on one lactation only is 0·25 in cows, and it is this value we shall use here. Values close to this have been obtained in many investigations, both in America and this country. The statement that the heritability of milk yield is one-quarter, put in non-statistical language, means that the difference in the breeding value of any two cows is, on the average, one-quarter of the difference between their records.

The expression for the estimated rate of genetic improvement for a known selection method has been given by Dickerson & Hazel (1944) in their paper on the progeny test. It may be desirable here to give a derivation of this expression as applied to our particular case.

The process of breeding and selection may be envisaged as follows. Young animals (heifers and young bulls) are taken into the breeding herd on the basis of their parents' performance and may themselves be regarded as the potential parents of the next generation. The decision as to which of these animals actually will pass their genes on is made on the basis of some measurement of their own genotype—by own performance in cows and by progeny tests in bulls. Genes may be transmitted to the next generation in four ways, which may be represented diagrammatically as follows:

A_1 Young animals, potential parents	Actual parents	Genetic superiority above same sex in A_1	Mean age when offspring born	A_2 Second generation, potential parents
Bulls	(1) Bulls to breed bulls	I_{BB}	L_{BB}	Bulls
	(2) Bulls to breed cows	I_{BC}	L_{BC}	
Cows	(3) Cows to breed bulls	I_{CB}	L_{CB}	Cows
	(4) Cows to breed cows	I_{CC}	L_{CC}	

Owing to the smaller number of males required, I_{CB} will necessarily be greater than I_{CC} and so on. In addition some of the classes of actual parents are mutually inclusive. For instance, a heifer calf from a cow of group 3 would be taken into the herd but not a bull calf from a cow of group 4.

The average rate of improvement may be obtained as follows. Let B and C be the mean genetic level of young animals born in a given year and taken into the breeding herd as potential parents. Then if ΔG is the genetic improvement per year, the mean genotype of sires of such bulls is $B - L_{BB}\Delta G + I_{BB}$ and of dams of such bulls $C - L_{CB}\Delta G + I_{CB}$, so that we have an equation

$$B = \tfrac{1}{2}(B - L_{BB}\Delta G + I_{BB} + C - L_{CB}\Delta G + I_{CB}),$$

and similarly

$$C = \tfrac{1}{2}(B - L_{BC}\Delta G + I_{BC} + C - L_{CC}\Delta G + I_{CC}).$$

On adding the two equations, B and C vanish and we are left with

$$\Delta G = \frac{\Sigma I}{\Sigma L}.$$

This gives the rate of genetic improvement to be expected from a selection programme in terms of the selection differentials and the generation lengths. In our case I_{BB} and I_{BC} are zero, since there is no selection amongst bulls.

The phenotype selection differential I_P measures the difference between the mean of the parents and the mean of all animals of the same sex in the population from which they were drawn. To convert this into a genetic measure, we must multiply by the heritability h^2. Then

$$I_G = I_P h^2.$$

If we assume that the phenotype is distributed normally, then we can express I_P as $\sigma_P \bar{\imath}$, where $\bar{\imath}$ is the superiority in terms of standard deviations and is obtainable in terms of the proportion selected, from tables (Pearson, 1931; Fisher & Yates, 1938), and σ_P is the phenotypic standard deviation. But also

$$\sigma_G = h\sigma_P,$$

where σ_G is the standard deviation of genotypic value.

Thus

$$I_G = \bar{\imath}h^2\sigma_P$$
$$= \bar{\imath}h\sigma_G.$$

The latter is the most convenient form, as for a given proportion selected the only variable is h.

Let us consider a herd of about 100 cows into which no breeding stock is introduced and in which there is no progeny testing and bulls are chosen solely on the milk records of their dam. The selection is being practised only in choosing cows to breed from. It is therefore necessary to calculate the possible selection differential in selecting cows to breed cows and selecting cows to breed bulls. As far as the former is concerned, let the proportion of cows in the first lactation be f_1, and so on, and let the natural mortality rates be 1/6 in the first three lactations and 1/3 for the fourth and after. Assume that we select on the first lactation only and reject all calves born from cows that are themselves culled. Let α be the number of mature heifers produced per cow in the herd per year.

Then the replacements available $= \alpha \ (f_2 + f_2 + f_3 + f_4 + \ldots)$, the first f_2 being calves born to cows, subsequently accepted into the herd, in the first lactation. Then

$$f_1 = \alpha \ (2f_2 + f_3 + f_4 + \ldots).$$

But we know from the mortality figures that at equilibrium

$$f_3 = \tfrac{5}{6} f_2, f_4 = \tfrac{5}{6} f_3, f_5 = \tfrac{2}{3} f_4,$$

so that the later proportions can be expressed in terms of f_2, giving

$$f_1 = \alpha \times 4 \cdot 916 f_2.$$

Of this f_1, one-sixth will die during the first lactation leaving $\tfrac{5}{6} \ (\alpha \times 4 \cdot 916 f_2)$, from which f_2 are to be selected.

The percentage culled is then $\dfrac{4 \cdot 097 \alpha - 1}{4 \cdot 097 \alpha} = 1 - \dfrac{0 \cdot 244}{\alpha}$. An accurate value of α is not available, but it may be assumed to lie between 0·35 and 0·40. This gives for percentage culled 39·0 and 30·3 respectively.

For milk yield, we may assume that $\sigma_P = 0 \cdot 20 \, \overline{Y}$ and $\sigma_G = 0 \cdot 10 \, \overline{Y}$, where $\overline{Y}$ is the average yield of the herd. Then the expected genetic superiority of the selected cows above their contemporaries will be $0 \cdot 032 \, \overline{Y}$ and $0 \cdot 024 \, \overline{Y}$ respectively. Calculation assuming culling at the end of the second lactation shows roughly the same results as on the first.

We may assume that the bulls are bred from the top 5 % of the herd (Lush, 1946, p. 147). The genetic superiority of their dams would then be $0 \cdot 103 \, \overline{Y}$. Assuming a sum of generation lengths of about 13 years, which seems to be about the minimum possible, we have for the expected genetic improvement

$$\Delta G = \frac{0 \cdot 135 \, \overline{Y}}{13} = 0 \cdot 010 \, \overline{Y} \quad \text{if} \quad \alpha = 0 \cdot 4$$

and

$$\Delta G = 0 \cdot 009 \, \overline{Y} \quad \text{for} \quad \alpha = 0 \cdot 35.$$

Thus, under the optimum conditions for mass selection the expected rate is 1·0 % per year. Of this, about one-quarter comes from the early culling of heifers and the rest from the selection of the dams of bulls.

The estimating of genetic improvement from the selection practised in a herd

Just as the rate of genetic improvement was estimated from the calculated superiority of parents over their contemporaries in the previous section, an estimation of the actual rate of genetic improvement in a herd can be obtained from the selection actually practised. Records of a herd founded in 1903 which is still in existence have been kindly made available to us. Although some breeding stock has been introduced during that time, about two-thirds of the animals born up to 1933 were sired by home-bred bulls. Thus it is possible to calculated the selection actually practised both in breeding cows and bulls. Selection was not all done at a given age as in our theoretical example, but there was a continual selection at all ages. The amount of selection was calculated as follows.

The records had, for other purposes, been divided into seven groups of sixty cows, the cows in each group being contemporaries. A calculation was therefore made within each group of the first lactation records of all cows having daughters in the milking herd, weighted according to the number of such daughters. This was compared with the average of all first lactations in the group. This gives the 'apparent' phenotypic selection differential as measured on first lactation and, on multiplying by 0·25, gives the 'apparent' genotypic

selection differential as shown in Table 1. This, however, does not represent the true selection differential if there has been selection in later lactations. Column 1 shows the actual amount of culling on first lactation as measured by the difference between the first lactations of animals which also had a second, and the whole group. The lack of correspondence between the two columns is a reflexion of the degree to which parenthood was independent of culling. Animals culled after their first lactation frequently left daughters in the herd and are therefore part of the selected breeding stock, whereas animals kept in the herd did not all have daughters. If all the selection was actually on the first lactation, then column 2 should be one-quarter of column 1. The fact that it is more than one-quarter indicates that some culling had also gone on in later lactations. The genetic improvement can then better be calculated on all the records available. The heritability then becomes dependent on the number of records, and the formula for genetic superiority is

$$I_{CC} = \Sigma n h^2{}_l (Y_l - \overline{Y}),$$

Table 1. *Estimation of selection differential from culling practised*

Period	Culling on 1st lactation* (gal.)	Genetic superiority of dams Based on 1st lact. (gal.)	Based on 1st–4th lact. (gal.)	No. of daughters in herd
1	+6·4	+6·1	+8·5	67
2	−7·2	+2·0	−3·5	44
3	+19·5	+3·5	+6·6	54
4	+2·7	+9·7	+16·2	30
5	+38·9	+0·7	+2·0	19
6	+34·0	+10·3	+17·6	24
7	+35·0	−1·4	−2·3	37
Total	+16·0	+4·8	+6·3	275

* Means of 1st lactations of cows having 2nd lactations – mean of all 1st lactations.

Table 2. *Selection differential on dams of bulls*

	Estimated genetic superiority of dams (based on 1st–4th lactations) (gal.)	No. of bulls	No. of daughters
Total	+38·1	67	289
Bull *AQ*	+101·4	1	65
Total less *AQ*	+17·5	66	224

where $h^2{}_l$ is heritability based on l lactations, $\overline{Y}$ is the mean of the first lactations of the whole group, Y_l is the mean of the lactations (corrected to first lactation) and n the number of daughters for each cow. Column 3 shows the values of I_{CC}, thus calculated, for each of the groups. The average value shows an increase of about one-third over that calculated on first lactations only. This value for I_{CC} should be compared with that theoretically possible as calculated above, which was, depending on the value of n, around 2·5-3 % of the herd average. This gives in this herd, where $\overline{Y}$ is 450 gal., a possible superiority of 12–14 gal. The fact that the actual figure is below this is due to many reasons —involuntary culling as, for example, for reactions to the tuberculin test, played some part, and in some stages selection was inefficient in the sense that a cow was culled after her first lactation, but nevertheless her heifer calf came into the herd.

In a similar way, the improvement due to selection of dams of bulls can be calculated. In this case each dam was weighted according to the number of daughters which her son had in the herd, as this varied considerably. This estimate is very uncertain, as one bull from an outstanding dam had sixty-five daughters in the herd. Exclusion of him from the

calculation reduces I_{CB} from 38 to 17. The results are shown in Table 2 and are calculated on all the available records of the dam. Using the figure for I_{CB} which includes the bull AQ, we have a value of $I_{CC} + I_{CB}$ of 44·1 gal.

The generation lengths were also calculated for the herd with the following results:

$$L_{BB} \quad 34 \text{ months}$$
$$L_{BC} \quad 36 \text{ months}$$
$$L_{CB} \quad 44 \text{ months}$$
$$L_{CC} \quad 47 \text{ months}$$

giving a total for ΣL of $13\frac{1}{2}$ years, and an estimate for the genetic improvement per year of $\Delta G = 44·1/13·5 = 3·3$ gal./year.

The generation interval is exceptionally short. Lush has estimated the mean generation interval (i.e. $\frac{1}{4}\Sigma L$) in dairy cattle at $4–4\frac{1}{2}$ years. An approximate check on British herd books gave an estimate of ΣL for Friesians and Shorthorns of 18·4 and 19·1 years respectively.

This estimate of the rate of genetic improvement of the herd has several snags. The first and most obvious one is in the applicability of heritability data derived from other herds and breeds of cattle to this particular herd. An estimate of heritability within the herd itself is in agreement with the figure used, but the standard error of the estimate is fairly large. Another source of error is in the sampling error of the estimate, both statistical and Mendelian. The value of the method lies not so much in the absolute value of the advance calculated as in the comparison of it with the optimum possible advance. In this case we have a value of I_{CC} of 6·3 gal. compared with a possible of 12–14 gal. and of I_{CB} of 38 gal. compared with a possible of 50 gal.

Discussion

The maximum possible rate of genetic improvement in a closed herd would seem to be of the order of 1 % of the average yield per year. In the herd whose records have been analysed, the probable improvement achieved by selection is 0·7 % per year, although the actual value could vary between fairly wide limits. The rate of improvement in normal pedigree herds would be expected to be less than this for several reasons:

(a) The rate of improvement depends on the mean generation length, which in the herd in question was just over 3 years. An examination of the Shorthorn and Friesian herd books shows that in those breeds the mean generation length was over $4\frac{1}{2}$ years. This would reduce the maximum improvement expected to about 0·6 % per year.

(b) Most herds are not closed and bulls are brought in. Knowledge of their breeding value would be less accurate than if they were home-bred, e.g. the management factor is not so well known.

(c) Selection will also be practised for other factors, e.g. type and butterfat percentage. This will reduce the selection differential on yield.

All these factors will reduce the rate of genetic improvement of breeds taken as a whole.

Practically all of the improvement expected comes from the selection of dams of bulls—only one-quarter comes from the selection of dams of cows. The contribution on the latter

side could be improved by anything which would increase the average number of heifer calves born in the lifetime of a cow. Amongst these would be:

(a) Increased longevity, although this would also increase the mean generation length.

(b) Sex determination, giving a larger selection differential in producing replacements.

(c) Superovulation and transplantation of ova of good cows into host cows with essentially the same effect as (b).

There is no immediate prospect of either (b) or (c) being applied commercially on a large scale.

It is often argued that the selection for productivity alone is liable to decrease the vigour and longevity of the animals produced. The evidence at present available on this point suggests that although there may be some such negative relationship between yield and longevity, it is not of very great effect. In addition, the cows are being naturally selected for longevity, in that after the first selection they are kept in the herd as long as they will live.

Progeny testing has been suggested as a method of increasing the rate of improvement. It has been shown (Dickerson & Hazel, 1944) that in a closed herd of 120 cows, it does not represent a useful method, as the increase in generation length more than offsets the selection differential. However, the use of artificial insemination in connexion with progeny testing makes it much more effective, and in a paper to be published shortly it will be shown that with a unit of size 2000 cows, it should be possible to obtain a rate of genetic improvement of about 1·8% of the average yield per year.

Summary

1. Methods of estimating the rate of genetic gain achieved in animal improvement are discussed and criticized.

2. A theoretical method of estimation of the maximum possible genetic gain of milk yield under direct selection is presented for a closed herd of dairy cattle. The maximum possible gain is approximately 1% of the average yield.

3. The method can also be used for the estimation of genetic gain from the selection practised in an actual herd. An estimated gain of milk yield of 0·7% per year is given for one herd.

REFERENCES

BONNIER, G., HANSSON, A. & SKJERVOLD, H. (1948). Studies on monozygous cattle twins. IX. The interplay of heredity and environment on growth and yield. *Acta Agric. Suec.* **3**, 1.

DICKERSON, G. E. & HAZEL, L. N. (1944). The effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. *J. Agric. Res.* **69**, 459.

FISHER, R. A. & YATES, F. (1938). *Statistical Tables for Biological, Agricultural and Medical Research.* Edinburgh: Oliver and Boyd.

LÖRTSCHER, H. (1937). Variationsstatistische Untersuchungen an Leistungs-Erhebungen in einer British Friesian Herde. *Z. Zücht.* B, **39**, 257.

LUSH, J. L. (1946). *Animal Breeding Plans.* Ames, Iowa: Iowa State College Press.

NELSON, R. H. (1943). Measuring the amount of genetic changes in a herd average. *J. Anim. Sci.* **2**, 358,

PEARSON, K. (Ed.) (1931). *Tables for Statisticians and Biometricians*, Part II. Cambridge University Press.